

SORT 30 (1) January-June 2006, 85-90

A matrix function useful in the estimation of linear continuous-time models

Heinz Neudecker

Cesaro

Abstract

In a recent publication Chen & Zadrozny (2001) derive some equations for efficiently computing e^A and ∇e^A , its derivative. They employ an expression due to Bellman (1960), Snider (1964) and Wilcox (1967) for the differential de^A and a method due to Van Loan (1978) to find the derivative ∇e^A . The present note gives *a)* a short derivation of ∇e^A by way of the Bellman-Snider-Wilcox result, *b)* a shorter derivation without using it. In both approaches there is no need for Van Loan's method.

MSC: 15A69

Keywords: matrix derivatives, vectorization, matrix exponential

1 Introduction

In a recent publication Chen & Zadrozny (2001) consider the matrix exponential

$$e^A = I_n + A + \frac{1}{2!}A^2 + \cdots + \frac{1}{k!}A^k + \cdots.$$

Their aim is to find ∇e^A which is *implicitly* defined by

$$d \operatorname{vec} e^A = (\nabla e^A) d \operatorname{vec} A$$

Address for correspondence: Oosterstraat, 13. 1741 GH Schagen. The Netherlands.

E-mail: H.Neudecker@uva.nl

Received: February 2005

Accepted: June 2006

or *explicitly* as

$$\nabla e^A = \frac{\partial \text{vec } e^A}{\partial (\text{vec } A)'}$$

They introduce a matrix function

$$K(t) = \int_{\tau=0}^t e^{(t-\tau)A'} \otimes e^{\tau A} d\tau$$

and recall that

$$d e^A = \int_{\tau=0}^1 e^{\tau A} (d A) e^{(1-\tau)A} d\tau.$$

See Bellman (1960, p. 171, (7)), Snider (1960) and Wilcox (1967). It then follows by vectorization that

$$d \text{vec } e^A = K(1) d \text{vec } A.$$

The authors compute $K(1)$ by employing a method due to Van Loan (1978). It turns out that $K(1)$ is the submatrix in the northeast corner of e^C , where

$$C = \begin{pmatrix} A' \otimes I_n & I_{n^2} \\ 0 & I_n \otimes A \end{pmatrix}.$$

The submatrix is denoted by $G_1(1)$. In this note we shall use two methods to find ∇e^A : *a)* one by way of the Bellman-Snider-Wilcox result, *b)* another more direct one without using that earlier result.

2 The derivation through $K(1)$

Starting from the expression

$$K(1) = \int_{\tau=0}^1 \left(e^{(1-\tau)A'} \otimes I_n \right) \left(I_n \otimes e^{\tau A} \right) d\tau,$$

which is also to be found in Chen & Zadrozny, we further develop

$$\begin{aligned} K(1) &= \int_{\tau=0}^1 e^{(1-\tau)A' \otimes I_n} e^{\tau I_n \otimes A} d\tau = \\ &= \int_{\tau=0}^1 e^{(1-\tau)A' \otimes I_n + \tau I_n \otimes A} d\tau = \end{aligned}$$

$$\begin{aligned}
&= \int_{\tau=0}^1 e^{A' \otimes I_n + \tau(I_n \otimes A - A' \otimes I_n)} d\tau = \\
&= e^{A' \otimes I_n} \int_{\tau=0}^1 e^{\tau(I_n \otimes A - A' \otimes I_n)} d\tau = \\
&= e^{A' \otimes I_n} \left[I_{n^2} + \frac{1}{2!} (I_n \otimes A - A' \otimes I_n) + \cdots + \frac{1}{(k+1)!} (I_n \otimes A - A' \otimes I_n)^k + \cdots \right] = \\
&= I_{n^2} + \frac{1}{2!} (I_n \otimes A + A' \otimes I_n) + \cdots + \frac{1}{(k+1)!} (I_n \otimes A + A' \otimes I_n)^{(k)} + \cdots
\end{aligned}$$

where for *commuting* A and B :

$$(A + B)^{(i)} = \sum_{j=1}^{i-1} A^j B^{i-j} + A^i + B^i, \quad (A + B)^{(1)} = A + B.$$

Clearly $I_n \otimes A$ and $A' \otimes I_n$ commute. We used Properties 3, 4 and 5 of the Appendix. It is clear that $K(1) = G_1(1)$, given the following computation. Consider

$$C = \begin{pmatrix} P & I \\ O & Q \end{pmatrix} \quad \text{and} \quad e^C = \begin{pmatrix} R & S \\ T & U \end{pmatrix}.$$

Then $T = O$, $R = I + \sum_{i=1}^{\infty} \frac{1}{i!} P^i = e^P$, $U = I + \sum_{i=1}^{\infty} \frac{1}{i!} Q^i = e^Q$ and

$$S = I_{n^2} + \sum_{i=1}^{\infty} \frac{1}{(i+1)!} (P + Q)^{(i)}.$$

Hence $G_1(1) = I_{n^2} + \sum_{i=1}^{\infty} \frac{1}{(i+1)!} (A' \otimes I + I \otimes A)^{(i)} = K(1)$. We can also define

$$C = \begin{pmatrix} I_n \otimes A & I_{n^2} \\ 0 & A' \otimes I_n \end{pmatrix}$$

to get the same $G_1(1)$.

3 A direct derivation of ∇e^A

Still simpler is to proceed as follows. Differentiation of e^A yields

$$d e^A = dA + \frac{1}{2!} \{(dA)A + AdA\} + \frac{1}{3!} \{(dA)A^2 + A(dA)A + A^2 dA\} + \dots$$

and from this by vectorization

$$\begin{aligned} d \operatorname{vec} e^A &= d \operatorname{vec} A + \frac{1}{2!} (I_n \otimes A + A' \otimes I_n) d \operatorname{vec} A + \frac{1}{3!} \{I_n \otimes A^2 + A' \otimes A + (A')^2 \otimes I\} \\ &\quad \times d \operatorname{vec} A + \dots = \\ &= \left[I_{n^2} + \frac{1}{2!} (I_n \otimes A + A' \otimes I_n) + \frac{1}{3!} (I_n \otimes A + A' \otimes I_n)^{(2)} + \dots \right] d \operatorname{vec} A = \\ &= \left[I_{n^2} + \sum_{i=1}^{\infty} \frac{1}{(i+1)!} (I_n \otimes A + A' \otimes I_n)^{(i)} \right] d \operatorname{vec} A. \end{aligned}$$

Hence

$$\nabla e^A = \frac{\partial \operatorname{vec} e^A}{\partial (\operatorname{vec} A)'} = I_{n^2} + \sum_{i=1}^{\infty} \frac{1}{(i+1)!} (I_n \otimes A + A' \otimes I)^{(i)}.$$

4 Appendix

Some algebraic properties:

1. $\operatorname{vec} ABC = (C' \otimes A) \operatorname{vec} B$.
2. $(A \otimes B)(C \otimes D) = AC \otimes BD$.
3. $e^{I \otimes A} = I \otimes e^A$, $e^{A \otimes I} = e^A \otimes I$.
4. $e^A \cdot e^B = e^{A+B}$ for commuting A and B .
5. $e^B \left[I + \frac{1}{2!}(A - B) + \frac{1}{3!}(A - B)^2 + \dots \right] = I + \frac{1}{2!}(A + B) + \frac{1}{3!}(A + B)^{(2)} + \dots$

$$\text{where } (A + B)^{(i)} = \sum_{j=1}^{i-1} A^j B^{i-j} + A^i + B^i \text{ for commuting } A \text{ and } B.$$

References

- Bellman, R. (1960). *Introduction to Matrix Analysis*. McGraw-Hill, New York.
- Chen, B. and Zadrozny, P. A. (2001). Analytic derivatives of the matrix exponential for estimation of linear continuous-time models. *Journal of Economic Dynamics & Control*, 25, 1867-1879.
- Neudecker, H. (1971). On a theorem of Snider and Wilcox. *METU Journal of Pure and Applied Sciences*, 4, 217-220.
- Snider, R. F. (1964). Perturbation variation methods for a quantum Boltzmann equation. *Journal of Mathematical Physics*, 5, 1580-1587.
- Van Loan, C. F. (1978). Computing integrals involving the matrix exponential. *IEEE Transactions on Automatic Control*, 23, 395-404.
- Wilcox, R. M. (1967). Exponential operators and parameter differentiation in quantum physics. *Journal of Mathematical Physics*, 8, 962-982.

